

RECEIVED

FEB 19 2003

TECH CENTER 1600/2900



1600

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/688,672ADATE: 02/11/2003  
TIME: 10:51:54Input Set : A:\-90-4-1.app  
Output Set: N:\CRF4\02112003\1688672A.raw

3 <110> APPLICANT: Skeiky, Yasir  
 4 Reed, Steven  
 5 Houghton, Raymond L.  
 6 McNeill, Patricia D.  
 7 Dillon, Davin C.  
 8 Lodes, Michael L.  
 9 Corixa Corporation  
 11 <120> TITLE OF INVENTION: Fusion Proteins of *Mycobacterium Tuberculosis*  
 13 <130> FILE REFERENCE: 014058-009041US  
 15 <140> CURRENT APPLICATION NUMBER: US 09/688,672A  
 16 <141> CURRENT FILING DATE: 2000-10-10  
 18 <150> PRIOR APPLICATION NUMBER: US 60/158,338  
 19 <151> PRIOR FILING DATE: 1999-10-07  
 21 <150> PRIOR APPLICATION NUMBER: US 60/158,425  
 22 <151> PRIOR FILING DATE: 1999-10-07  
 24 <160> NUMBER OF SEQ ID NOS: 202  
 26 <170> SOFTWARE: PatentIn Ver. 2.1  
 28 <210> SEQ ID NO: 1  
 29 <211> LENGTH: 2220  
 30 <212> TYPE: DNA  
 31 <213> ORGANISM: *Mycobacterium tuberculosis*  
 33 <220> FEATURE:  
 34 <221> NAME/KEY: CDS  
 35 <222> LOCATION: (1)..(2220)  
 36 <223> OTHER INFORMATION: Mtb81  
 38 <400> SEQUENCE: 1  
 39 act gat cgc gtg ttg ggc aac 10 48  
 40 Thr Asp Arg Val Ser Val Gly Asn Leu Arg Ile Ala Arg Val Leu Tyr 15  
 41 1 5 48  
 43 gac ttc gtg aac aat gaa 25 96  
 44 Asp Phe Val Asn Asn Glu Ala Leu Pro Gly Thr Asp Ile Asp Pro Asp 10  
 45 20 30 144  
 47 agc ttc tgg gcg gtc gac aar 40 192  
 48 Ser Phe Trp Ala Gly Val Asp Lys Val Val Ala Asp Leu Thr Pro Gln 45  
 49 35 60 240  
 51 aac caa gct ctg ttg aac 55 80 288  
 52 Asn Gln Ala Leu Leu Asn Ala Arg Asp Glu Leu Gln Ala Gln Ile Asp 75  
 53 50 70  
 55 aag tgg cac cgg cgt cgg 70 80  
 56 Lys Trp His Arg Arg Arg Val Ile Glu Pro Ile Asp Met Asp Ala Tyr 60  
 59 cgc cag ttc ctc acc gag atc 70  
 60 Arg Gln Phe Leu Thr Glu Ile Gly Tyr Leu Leu Pro Glu Pro Asp Asp 60

P.6

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/688,672A

DATE: 02/11/2003  
TIME: 10:51:54

Input Set : A:\-90-4-1.app  
Output Set: N:\CRF4\02112003\I688672A.raw

61	85	90	95	
63	ttc acc atc acc acg tcc ggt gtc gac gct gag atc acc acg acc gcc			336
64	Phe Thr Ile Thr Thr Ser Gly Val Asp Ala Glu Ile Thr Thr Ala			
65	100	105	110	
67	ggc ccc cag ctg gtg gtg ccg ctc aac gcg cgg ttt gct ctg aac			384
68	Gly Pro Gln Leu Val Val Pro Val Leu Asn Ala Arg Phe Ala Leu Asn			
69	115	120	125	
71	gct gac aac gct cgc tgg ggc tcc ctc tac gac gcc ttg tat ggc acc			432
72	Ala Ala Asn Ala Arg Trp Gly Ser Leu Tyr Asp Ala Leu Tyr Gly Thr			
73	130	135	140	
75	gat gtc atc ccc gag acc gac ggc gcc gaa aaa ggc ccc acg tac aac			480
76	Asp Val Ile Pro Glu Thr Asp Gly Ala Glu Lys Gly Pro Thr Tyr Asn			
77	145	150	155	160
79	aag gtt cgt ggc gac aag gtg atc gct tat gcc cgc aag ttc ctc gac			528
80	Lys Val Arg Gly Asp Lys Val Ile Ala Tyr Ala Arg Lys Phe Leu Asp			
81	165	170	175	
83	gac agt gtt ccg ctg tcg tcg ggt tcc ttt ggc gac gcc acc ggt ttc			576
84	Asp Ser Val Pro Leu Ser Ser Gly Ser Phe Gly Asp Ala Thr Gly Phe			
85	180	185	190	
87	aca gtg cag gat ggc cag ctc gtg gtt gcc ttg ccg gat aag tcc acc			624
88	Thr Val Gln Asp Gly Gln Leu Val Val Ala Leu Pro Asp Lys Ser Thr			
89	195	200	205	
91	ggc ctg gcc aac ccc ggc cag ttc gcc ggc tac acc ggc gca gcc gag			672
92	Gly Leu Ala Asn Pro Gly Gln Phe Ala Gly Tyr Thr Gly Ala Ala Glu			
93	210	215	220	
95	tcg ccg aca tcg ctg cta atc aat cac ggt ttg cac atc gag atc			720
96	Ser Pro Thr Ser Val Leu Leu Ile Asn His Gly Leu His Ile Glu Ile			
97	225	230	235	240
99	ctg atc gat ccg gag tcg cag gtc ggc acc acc gac cgg gcc ggc gtc			768
100	Leu Ile Asp Pro Glu Ser Gln Val Gly Thr Thr Asp Arg Ala Gly Val			
101	245	250	255	
103	aag gac gtg atc ctg gaa tcc gcg atc acc acg atc atg gac ttc gag			816
104	Lys Asp Val Ile Leu Glu Ser Ala Ile Thr Thr Ile Met Asp Phe Glu			
105	260	265	270	
107	gac tcg gtg gcc gtc gac gcc gac aag gtg ctg ggt tat cgg			864
108	Asp Ser Val Ala Ala Val Asp Ala Ala Asp Lys Val Leu Gly Tyr Arg			
109	275	280	285	
111	aac tgg ctc ggc ctg aac aag ggc gac ctg gca gca gcg gta gac aag			912
112	Asn Trp Leu Gly Leu Asn Lys Gly Asp Leu Ala Ala Val Asp Lys			
113	290	295	300	
115	gac ggc acc gct ttc ctg cgg gtg ctc aat agg gac cgg aac tac acc			960
116	Asp Gly Thr Ala Phe Leu Arg Val Leu Asn Arg Asp Arg Asn Tyr Thr			
117	305	310	315	320
119	gca ccc ggc ggt ggc cag ttc acg ctg cct gga cgc agc ctc atg ttc			1008
120	Ala Pro Gly Gly Gln Phe Thr Leu Pro Gly Arg Ser Leu Met Phe			
121	325	330	335	
123	gtc cgc aac gtc ggt cac ttg atg acg aat gac gcc atc gtc gac act			1056
124	Val Arg Asn Val Gly His Leu Met Thr Asn Asp Ala Ile Val Asp Thr			
125	340	345	350	

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/688,672A**

DATE: 02/11/2003  
 TIME: 10:51:54

Input Set : A:\-90-4-1.app  
 Output Set: N:\CRF4\02112003\I688672A.raw

127	gac	ggc	agc	gag	gtg	ttc	gaa	ggc	atc	atg	gat	gcc	cta	ttc	acc	ggc	1104
128	Asp	Gly	Ser	Glu	Val	Phe	Glu	Gly	Ile	Met	Asp	Ala	Leu	Phe	Thr	Gly	
129									355				360			365	
131	ctg	atc	gcc	atc	cac	ggg	cta	aag	gcc	agc	gac	gtc	aac	ggg	ccg	ctg	1152
132	Leu	Ile	Ala	Ile	His	Gly	Leu	Lys	Ala	Ser	Asp	Val	Asn	Gly	Pro	Leu	
133									370				375			380	
135	atc	aac	agc	cgc	acc	ggc	tcc	atc	tac	atc	gtc	aag	ccg	aag	atg	cac	1200
136	Ile	Asn	Ser	Arg	Thr	Gly	Ser	Ile	Tyr	Ile	Val	Lys	Pro	Lys	Met	His	
137									385				390			395	400
139	ggt	ccg	gcc	gag	gtg	gcg	ttt	acc	tgc	gaa	ctg	ttc	agc	ccg	gtt	gaa	1248
140	Gly	Pro	Ala	Glu	Val	Ala	Phe	Thr	Cys	Glu	Leu	Phe	Ser	Arg	Val	Glu	
141									405				410			415	
143	gat	gtg	ctg	ggg	ttg	ccg	caa	aac	acc	atg	aag	atc	ggc	atc	atg	gac	1296
144	Asp	Val	Leu	Gly	Leu	Pro	Gln	Asn	Thr	Met	Lys	Ile	Gly	Ile	Met	Asp	
145									420				425			430	
147	gag	gaa	cgc	ccg	acc	acg	gtc	aac	ctc	aag	gcc	tgc	atc	aaa	gct	gcc	1344
148	Glu	Glu	Arg	Arg	Thr	Thr	Val	Asn	Leu	Lys	Ala	Cys	Ile	Lys	Ala	Ala	
149									435				440			445	
151	gcg	gac	cgc	gtg	gtg	ttc	atc	aac	acc	ggg	ttc	ctg	gac	cgc	acc	ggc	1392
152	Ala	Asp	Arg	Val	Val	Phe	Ile	Asn	Thr	Gly	Phe	Leu	Asp	Arg	Thr	Gly	
153									450				455			460	
155	gat	gaa	atc	cac	acc	tcg	atg	gag	gcc	ccg	atg	gtg	cgc	aag	ggc	1440	
156	Asp	Glu	Ile	His	Thr	Ser	Met	Glu	Ala	Gly	Pro	Met	Val	Arg	Lys	Gly	
157									465				470			475	480
159	acc	atg	aag	agc	cag	ccg	tgg	atc	ttg	gcc	tac	gag	gac	cac	aac	gtc	1488
160	Thr	Met	Lys	Ser	Gln	Pro	Trp	Ile	Leu	Ala	Tyr	Glu	Asp	His	Asn	Val	
161									485				490			495	
163	gat	gcc	ggc	ctg	gcc	ggc	ggg	ttc	agc	ggc	cga	gcc	cag	gtc	ggc	aag	1536
164	Asp	Ala	Gly	Leu	Ala	Ala	Gly	Phe	Ser	Gly	Arg	Ala	Gln	Val	Gly	Lys	
165									500				505			510	
167	ggc	atg	tgg	aca	atg	acc	gag	ctg	atg	gcc	gac	atg	gtc	gag	aca	aaa	1584
168	Gly	Met	Trp	Thr	Met	Thr	Glu	Leu	Met	Ala	Asp	Met	Val	Glu	Thr	Lys	
169									515				520			525	
171	atc	gcc	cag	ccg	cgc	gcc	ggg	gcc	agc	acc	gcc	tgg	gtt	ccc	tct	ccc	1632
172	Ile	Ala	Gln	Pro	Arg	Ala	Gly	Ala	Ser	Thr	Ala	Trp	Val	Pro	Ser	Pro	
173									530				535			540	
175	act	gcg	gcc	acc	ctg	cat	gcg	ctg	cac	tac	cac	cag	gtc	gac	gtc	gcc	1680
176	Thr	Ala	Ala	Thr	Leu	His	Ala	Leu	His	Tyr	His	Gln	Val	Asp	Val	Ala	
177									545				550			555	560
179	gcg	gtg	caa	caa	gga	ctg	gcg	ggg	aag	cgt	cgc	gcc	acc	atc	gaa	caa	1728
180	Ala	Val	Gln	Gln	Gly	Leu	Ala	Gly	Lys	Arg	Arg	Ala	Thr	Ile	Glu	Gln	
181									565				570			575	
183	ttg	ctg	acc	att	ccg	ctg	gcc	aag	gaa	ttg	gcc	tgg	gtc	ccc	gac	gag	1776
184	Leu	Leu	Thr	Ile	Pro	Leu	Ala	Lys	Glu	Leu	Ala	Trp	Ala	Pro	Asp	Glu	
185									580				585			590	
187	atc	cgc	gaa	gag	gtc	gac	aac	aac	tgt	caa	tcc	atc	ctc	ggc	tac	gtg	1824
188	Ile	Arg	Glu	Glu	Val	Asp	Asn	Asn	Cys	Gln	Ser	Ile	Leu	Gly	Tyr	Val	
189									595				600			605	
191	gtt	cgc	tgg	gtt	gat	caa	ggt	gtc	ggc	tgc	tgc	aag	gtg	ccc	gac	atc	1872

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/688,672A

DATE: 02/11/2003  
 TIME: 10:51:54

Input Set : A:\-90-4-1.app  
 Output Set: N:\CRF4\02112003\I688672A.raw

192 Val Arg Trp Val Asp Gln Gly Val Gly Cys Ser Lys Val Pro Asp Ile			
193 610	615	620	
195 cac gac gtc gcg ctc atg gag gac cgg gcc acg ctg cga atc tcc agc			1920
196 His Asp Val Ala Leu Met Glu Asp Arg Ala Thr Leu Arg Ile Ser Ser			
197 625	630	635	640
199 caa ttg ttg gcc aac tgg ctg cgc cac ggt gtt atc acc agc gcg gat			
200 Gln Leu Leu Ala Asn Trp Leu Arg His Gly Val Ile Thr Ser Ala Asp			1968
201 645	650	655	
203 gtt cgg gcc agc ttg gag cgg atg gct cgg ttg gtc gat cga caa aac			2016
204 Val Arg Ala Ser Leu Glu Arg Met Ala Pro Leu Val Asp Arg Gln Asn			
205 660	665	670	
207 gcg ggc gac gtt gca tac cga ccg atg gca ccc aac ttc gac gac agt			2064
208 Ala Gly Asp Val Ala Tyr Arg Pro Met Ala Pro Asn Phe Asp Asp Ser			
209 675	680	685	
211 atc gcc ttc ctg gcc cag gag ctg atc ttg tcc ggg gcc cag cag			2112
212 Ile Ala Phe Leu Ala Ala Gln Glu Leu Ile Leu Ser Gly Ala Gln Gln			
213 690	695	700	
215 ccc aac ggc tac acc gag ccg atc ctg cac cga cgt cgt cgg gag ttt			2160
216 Pro Asn Gly Tyr Thr Glu Pro Ile Leu His Arg Arg Arg Arg Glu Phe			
217 705	710	715	720
219 aag gcc cgg gcc gct gag aag ccg gcc cca tcg gac agg gcc ggt gac			
220 Lys Ala Arg Ala Ala Glu Lys Pro Ala Pro Ser Asp Arg Ala Gly Asp			2208
221 725	730	735	
223 gat gcg gcc agg			
224 Asp Ala Ala Arg			2220
225 740			
228 <210> SEQ ID NO: 2			
229 <211> LENGTH: 740			
230 <212> TYPE: PRT			
231 <213> ORGANISM: Mycobacterium tuberculosis			
233 <400> SEQUENCE: 2			
234 Thr Asp Arg Val Ser Val Gly Asn Leu Arg Ile Ala Arg Val Leu Tyr			
235 1 5 10 15			
237 Asp Phe Val Asn Asn Glu Ala Leu Pro Gly Thr Asp Ile Asp Pro Asp			
238 20 25 30			
240 Ser Phe Trp Ala Gly Val Asp Lys Val Val Ala Asp Leu Thr Pro Gln			
241 35 40 45			
243 Asn Gln Ala Leu Leu Asn Ala Arg Asp Glu Leu Gln Ala Gln Ile Asp			
244 50 55 60			
246 Lys Trp His Arg Arg Arg Val Ile Glu Pro Ile Asp Met Asp Ala Tyr			
247 65 70 75 80			
249 Arg Gln Phe Leu Thr Glu Ile Gly Tyr Leu Leu Pro Glu Pro Asp Asp			
250 85 90 95			
252 Phe Thr Ile Thr Thr Ser Gly Val Asp Ala Glu Ile Thr Thr Ala			
253 100 105 110			
255 Gly Pro Gln Leu Val Val Pro Val Leu Asn Ala Arg Phe Ala Leu Asn			
256 115 120 125			
258 Ala Ala Asn Ala Arg Trp Gly Ser Leu Tyr Asp Ala Leu Tyr Gly Thr			
259 130 135 140			

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/688,672A

DATE: 02/11/2003

TIME: 10:51:54

Input Set : A:\-90-4-1.app

Output Set: N:\CRF4\02112003\I688672A.raw

261 Asp Val Ile Pro Glu Thr Asp Gly Ala Glu Lys Gly Pro Thr Tyr Asn  
 262 145 150 155 160  
 264 Lys Val Arg Gly Asp Lys Val Ile Ala Tyr Ala Arg Lys Phe Leu Asp  
 265 165 170 175  
 267 Asp Ser Val Pro Leu Ser Ser Gly Ser Phe Gly Asp Ala Thr Gly Phe  
 268 180 185 190  
 270 Thr Val Gln Asp Gly Gln Leu Val Val Ala Leu Pro Asp Lys Ser Thr  
 271 195 200 205  
 273 Gly Leu Ala Asn Pro Gly Gln Phe Ala Gly Tyr Thr Gly Ala Ala Glu  
 274 210 215 220  
 276 Ser Pro Thr Ser Val Leu Leu Ile Asn His Gly Leu His Ile Glu Ile  
 277 225 230 235 240  
 279 Leu Ile Asp Pro Glu Ser Gln Val Gly Thr Thr Asp Arg Ala Gly Val  
 280 245 250 255  
 282 Lys Asp Val Ile Leu Glu Ser Ala Ile Thr Thr Ile Met Asp Phe Glu  
 283 260 265 270  
 285 Asp Ser Val Ala Ala Val Asp Ala Ala Asp Lys Val Leu Gly Tyr Arg  
 286 275 280 285  
 288 Asn Trp Leu Gly Leu Asn Lys Gly Asp Leu Ala Ala Val Asp Lys  
 289 290 295 300  
 291 Asp Gly Thr Ala Phe Leu Arg Val Leu Asn Arg Asp Arg Asn Tyr Thr  
 292 305 310 315 320  
 294 Ala Pro Gly Gly Gln Phe Thr Leu Pro Gly Arg Ser Leu Met Phe  
 295 325 330 335  
 297 Val Arg Asn Val Gly His Leu Met Thr Asn Asp Ala Ile Val Asp Thr  
 298 340 345 350  
 300 Asp Gly Ser Glu Val Phe Glu Gly Ile Met Asp Ala Leu Phe Thr Gly  
 301 355 360 365  
 303 Leu Ile Ala Ile His Gly Leu Lys Ala Ser Asp Val Asn Gly Pro Leu  
 304 370 375 380  
 306 Ile Asn Ser Arg Thr Gly Ser Ile Tyr Ile Val Lys Pro Lys Met His  
 307 385 390 395 400  
 309 Gly Pro Ala Glu Val Ala Phe Thr Cys Glu Leu Phe Ser Arg Val Glu  
 310 405 410 415  
 312 Asp Val Leu Gly Leu Pro Gln Asn Thr Met Lys Ile Gly Ile Met Asp  
 313 420 425 430  
 315 Glu Glu Arg Arg Thr Thr Val Asn Leu Lys Ala Cys Ile Lys Ala Ala  
 316 435 440 445  
 318 Ala Asp Arg Val Val Phe Ile Asn Thr Gly Phe Leu Asp Arg Thr Gly  
 319 450 455 460  
 321 Asp Glu Ile His Thr Ser Met Glu Ala Gly Pro Met Val Arg Lys Gly  
 322 465 470 475 480  
 324 Thr Met Lys Ser Gln Pro Trp Ile Leu Ala Tyr Glu Asp His Asn Val  
 325 485 490 495  
 327 Asp Ala Gly Leu Ala Ala Gly Phe Ser Gly Arg Ala Gln Val Gly Lys  
 328 500 505 510  
 330 Gly Met Trp Thr Met Thr Glu Leu Met Ala Asp Met Val Glu Thr Lys  
 331 515 520 525  
 333 Ile Ala Gln Pro Arg Ala Gly Ala Ser Thr Ala Trp Val Pro Ser Pro

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/688,672A

DATE: 02/11/2003  
TIME: 10:51:55

Input Set : A:\-90-4-1.app  
Output Set: N:\CRF4\02112003\I688672A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 406  
Seq#:11; N Pos. 325,328,348,353,442,659,679  
Seq#:12; Xaa Pos. 63,121,285  
Seq#:29; N Pos. 1460,1854  
Seq#:33; N Pos. 497,500,1136,1445,1487,1509,1515  
Seq#:155; N Pos. 104